Insect Molecular Biology (2000) **9**(1), 19–26

cDNA sequence, mRNA expression and genomic DNA of trypsinogen from the Indianmeal moth, *Plodia interpunctella*

Y. C. Zhu, B. Oppert, K. J. Kramer, W. H. McGaughey and A. K. Dowdy

Grain Marketing and Production Research Center, Agricultural Research Service, US Department of Agriculture, Manhattan, Kansas, USA

Abstract

Trypsin-like enzymes are major insect gut enzymes that digest dietary proteins and proteolytically activate insecticidal proteins produced by the bacterium *Bacillus thuringiensis* (Bt). Resistance to Bt in a strain of the Indianmeal moth, *Plodia interpunctella*, was linked to the absence of a major trypsin-like proteinase (Oppert *et al.*, 1997). In this study, trypsin-like proteinases, cDNA sequences, mRNA expression levels and genomic DNAs from Bt-susceptible and -resistant strains of the Indianmeal moth were compared. Proteinase activity blots of gut extracts indicated that the susceptible strain had two major trypsin-like proteinases, whereas the resistant strain had only one. Several trypsinogen-like cDNA clones were isolated and sequenced from cDNA libraries of both strains using a probe deduced from a conserved sequence for a serine proteinase active site. cDNAs of 852 nucleotides from the susceptible strain and 848 nucleotides from the resistant strain contained an open reading frame of 783 nucleotides which encoded a 261-amino acid trypsinogen-like protein. There was a single silent nucleotide difference between the two cDNAs in the open reading frame and the predicted amino acid sequence from the cDNA clones was most similar to sequences of trypsin-like proteinases from the spruce budworm, *Choristoneura fumiferana*, and the tobacco hornworm, *Manduca sexta*. The encoded protein included amino acid sequence motifs of serine proteinase active sites, conserved cysteine residues, and both zymogen activation and signal peptides. Northern blotting analysis showed no major difference between the two strains in mRNA expression in fourth-instar larvae, indicating that transcription was similar in the strains. Southern blotting analysis revealed that the restriction sites for the trypsinogen genes from the susceptible and resistant strains were different. Based on an enzyme size comparison, the cDNA isolated in this study corresponded to the gene for the smaller of two trypsin-like proteinases, which is found in both the Bt-susceptible and -resistant strains of the Indianmeal moth. The sequences reported in this paper have been deposited in the GenBank database (accession numbers AF064525 for the RC688 strain and AF064526 for HD198).

Keywords: trypsinogen, *Plodia interpunctella*, *Bacillus thuringiensis*, cDNA, gene, amino acid sequence, Indianmeal moth, gut, proteinase.

Received 17 September 1998; accepted 8 December 1998. Correspondence: A. K. Dowdy, Grain Marketing and Production Research Center, Agricultural Research Service, US Department of Agriculture, 1515 College Avenue, Manhattan, Kansas 66502–2736, USA. e-mail: dowdy@usgmrl.ksu.edu

Introduction

The Indianmeal moth, *Plodia interpunctella* (Hübner), is an economically important stored-product insect. Like most lepidopterans, larvae of the Indianmeal moth possess midguts with an alkaline pH, and the major digestive enzymes are serine proteinases (Oppert *et al.*, 1996). Trypsins are serine proteinases that cleave intact protein chains on the carboxyl side of basic L-amino acids such as arginine or lysine (Walsh & Wilcox, 1970). In addition to hydrolysing dietary protein, trypsin-like enzymes are involved in the activation of protoxins from *Bacillus thuringiensis* (Bt) (Milne & Kaplan, 1993; Oppert *et al.*, 1996). To bind target tissue receptors, the protoxin must be solubilized and proteolytically cleaved to an activated form (Ogiwara *et al.*, 1992). Recent evidence indicates that the lack of trypsin-like gut proteinases enables some insects to adapt to Bt toxins via a mechanism where incomplete or non-activation of the protoxin occurs (Oppert *et al.*, 1997).

To understand how insects regulate gut proteinase

activity and to provide fundamental information for Bt resistance management programmes, we investigated the midgut trypsin-like proteinases and their genes in Bt-susceptible (RC688) and -resistant (HD198) strains of *P. interpunctella* using proteinase, cDNA, and Northern and Southern blot analyses. Trypsin-like enzyme activity, cDNA sequences, predicted amino acid sequences, levels of mRNA expression, and genomic DNA structures for a trypsinogen-like protein of the moth are reported.

Results

Midgut proteinase activity

Phenotypic expression of trypsin-like proteinases varied between the two strains of the Indianmeal moth (Fig. 1). Gut extracts from RC688 had two major BApNA-hydrolysing proteinase activities, labelled T1 and T2. The molecular masses of these proteinases are approximately 45 and 25 kDa, respectively. Strain HD198, however, only expresses an active T2 proteinase. This result agreed with previous observations, in which the lack of proteinase T1 was genetically linked to survival of *P. interpunctella* on Bt-treated diets (Oppert *et al.*, 1997).

cDNA sequences

cDNA libraries from Indianmeal moth strains RC688 and HD198 were screened using a trypsinogen-like enzyme cDNA probe. We successfully cloned the cDNA for the T2 proteinase that does not appear to be involved in activation of the Bt protoxin in the midgut of susceptible moths. Preliminary sequencing identified three clones from each cDNA library which had 3′ ends (≈ 300 bp) identical to the trypsinogen-like cDNA sequence. Two of the clones were selected from each library for sequencing and characterization.

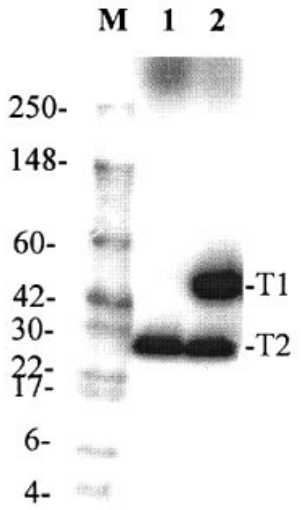

Figure 1. Hydrolysis of BApNA by larval gut proteinases from Indianmeal moth strain HD198 (lane 1) and RC688 (lane 2). M, molecular mass markers.

Two identical cDNAs of a putative trypsinogen-like protein were obtained from the RC688 library, which contained 852 nucleotides and had the start codon ATG at positions 1–3 and the termination codon TAA at positions 784–786 (Fig. 2). Two identical clones also were obtained from the HD198 library, but these contained only 848 nucleotides. The open reading frame for all four clones consisted of 783 nucleotides which encoded for 261 amino acid residues and corresponded to an immature form of a trypsin-like proteinase. The polyadenylation signal, AATAAA, was located at positions 803–808 in both of the cDNAs. Alignment of the cDNA sequences revealed that only one nucleotide was different between the RC688 and HD198 strains. A nucleotide thymine at position 738 in the RC688 cDNA was substituted with a cytosine in the HD198 cDNA. However, this change did not alter the encoded amino acid, valine.

Sequence comparison

A search of the GenBank database using the Blastx non-redundant program revealed that the sequence of the *P. interpunctella* protein was similar to proteins in the serine proteinase trypsin/chymotrypsin family. The most similar sequences included trypsin-like enzymes from the spruce budworm, *Choristoneura fumiferana* (Clemens) (Wang *et al.*, 1993), tobacco hornworm, *Manduca sexta* (L.) (Peterson *et al.*, 1994), sheep blowfly, *Lucilia cuprina* (Casu *et al.*, 1994), malaria mosquito, *Anopheles gambiae* (Muller *et al.*, 1993) and a serine protease from the common cattle grub, *Hypoderma lineatum* (Villers) (Moire *et al.*, 1994). The mature trypsin of *P. interpunctella* was most similar to the trypsin-like enzyme of *C. fumiferana* with 75% similarity and 68% identity in the amino acid residues (Table 1). Sequence analyses using GCG Distances and Growtree methods also indicated that the trypsinogen-like protein from *P. interpunctella* was most similar to the trypsin-like proteinases from *C. fumiferana* and *M. sexta* (Table 1).

The predicted amino acid sequence encoded by the *P. interpunctella* cDNA was aligned with five homologous insect proteases (Fig. 3). This sequence contained all of the conserved residues in the putative active site, His^{74}, Asp^{120} and Ser^{218}, which form the catalytic triad in serine proteases (Wang *et al.*, 1993; Peterson *et al.*, 1994). Six cysteine residues, predicted to occur in disulphide bridge configurations among trypsins and chymotrypsins, were located at positions 58, 75, 185, 202, 214 and 238. The same pattern of cysteine residues was found in all five insect serine proteases. A seventh cysteine also was present in the three lepidopteran species, *P. interpunctella* (position 232), *C. fumiferana* and *M. sexta*. The deduced amino acid sequences of the trypsinogen-like proteinases from both the RC688 and HD198 *P. interpunctella* cDNAs contained a putative signal peptide fifteen amino acid

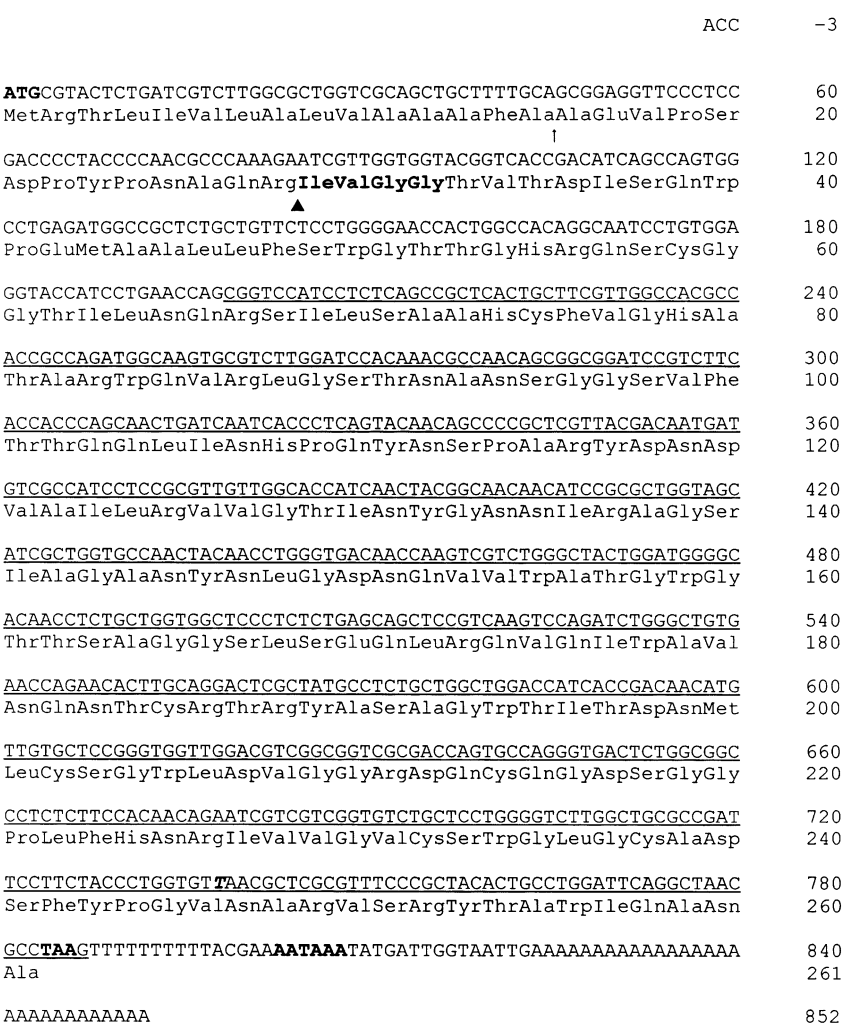

Figure 2. Nucleotide and deduced amino acid sequences of trypsinogen-like cDNA isolated from *P. interpunctella* gut RNA. ATG, start codon; TAA, termination codon; AATAAA, polyadenylation signal; ↑, predicted signal peptide cleavage site; ▲, predicted activation peptide cleavage site. IleValGlyGly are conserved N-terminal residues. Sequence identical to gut trypsinogen-like cDNA sequence (obtained from gut tissue using RT-PCR) is underlined. A thymine at position 738 (*T*) in the RC688 cDNAs was substituted for a cytosine in the HD198 cDNAs.

Table 1. Similarity comparison and evolutionary distance relationship (GCG) of a *P. interpunctella* trypsin-like proteinase to serine proteinases from five other insect species.

Species	Proteinase Type	Similarity (%)	Identity (%)	Evolutionary Distance
P. interpunctella	Trypsin	—	—	
C. fumiferana	Trypsin	75	68	
M. sexta	Trypsin	72	63	
L. cuprina	Trypsin	55	40	
H. lineatum	Serine proteinase	55	34	
A. gambiae	Trypsin	53	36	

residues long and an activation peptide thirteen amino acid residues long. The calculated molecular masses for the precursor of trypsinogen and the mature enzyme were 27,927 and 24,959 Da, respectively. The sequence IVGG at positions 29–32 was highly conserved in many trypsin- and chymotrypsin-like proteinases and marked the N-termini of the active enzymes (Wang *et al.*, 1993).

Trypsinogen-like cDNA sequence from gut tissues

A 589 bp clone, which contained a sequence encoding the serine in the active site, was obtained using reverse transcription polymerase chain reaction (RT-PCR) amplification of gut RNA from the RC688 strain. Sequence alignment revealed that the cDNA sequence in gut tissues was identical to the cDNA sequence obtained from the whole-body cDNA library (Fig. 2). These data indicated that the cDNA isolated from the whole-body cDNA library was the one that codes for a gut trypsinogen-like proteinase which was most similar to that of *C. fumiferana* (Wang *et al.*, 1995).

Trypsin-like gene expression in two P. interpunctella *strains*

To examine whether trypsin-like enzyme gene expression was the same in the two *P. interpunctella* strains, Northern blots containing mRNAs from RC688 and HD198 were hybridized with a PCR-generated probe. The data showed that the trypsinogen-like cDNA probe hybridized to mRNAs

```
      1                     ↓                          ⇓                        75
PiTP  .M..RTLIVL....ALVAAAFAAE.......V...PSDP....YPNAQRIVGGTVTDISQWPEMAALLFSWGTTG
CfTP  .M..R..VTL....ALVALCLVS........V...AALP.....EKQQRIVGGSVTTIEQWPSGSALLYSWNLVT
MsTP  .M..R..LFL....ALLALGFAA........V...AAVP.....ANPQRIVGGSTTTIQQYPTIVALLFSRNGNT
LcTP  ML..KFVVLLC...A.ISCALGA.......AV.PEGMVPQ.....LDGRIVGGVATTISSFPWQISL..QRSG..
HlSP  ML..KFVILLC...S.IAYVFGA.......VV.PLGMLSQ.....SDGRIVGGVESKIEDFPWQISL..QRDG..
AgTP  .MSNKIAILLAVLVAVVACA.EAQANQRHRLVRPSPSFSPRPRYAVGQRIVGGFEIDVSDAPYQVSL..QYNK..
               *                     *                *****        *    *

      76  ◆               ▼◆                                                   ▼
PiTP  HRQSCGGTILNQRSI.LSAAHCFVGHATARWQVRLGST.NANSGGSVFTTQQL....INHPQY.NSPAR.YDNDV
CfTP  YSQACGGAILNTRSI.LSAAHCFIGDAANRWRIRTGST.WANSGGVVHNT.AL..III.HPSY.NT..RTLDNDI
MsTP  FFQACGGTILNNRNV.LTAAHCPHGDAVNRWRVRSGST.FANSGGAVHN...LNSVRI.HPNY.NR..RNLDNDI
LcTP  .SHSCGGSVYNSR.IIVTAAHCLQSVSTSVLKVRRGSSYWN.SGGVVVSVAAF....KNHEGY.NP..KTMVNDI
HlSP  .RHYCGGSIY.SKNVIITAAHCLRNVVAEELRVRVGSSYWE.HGGSLRNISKF....QIHESYVEP...TKEYDD
AgTP  .RHNCGGSVLSSKWV.LTAAHCTAGASTSSLTVRLGTS.RHASGGTVVRVARV....VQHPKY.DS..SSIDFDY
          ***           ****           * *       **              *  *          *

      151                                                                ◆    225
PiTP  AILRVVGTINYGNNIRAGSIAGANYNLGDNQVVWAT..GWGTT..SAGGSLSEQLRQVQIWAVNQNTCRTR.YAS
CfTP  AILRSATTIAQNNQARPASIAGANYNLADNQAVWAI..GWGAT..CPGCAGSEQLRHVQIWTVNQNTCRSR.YLE
MsTP  AIMRTASNIAFNNAAQPARIAGANYNLGDNQVVWAA..GWGAI..RSGGPSSEQLRHVQVWTVNQATCRSR.YAS
LcTP  AVIRLSSSLTMSSTIKAIALT..T.AAPANGAA.ATVSGWGTT..SSGGSIPAQLRYVDLKIVGRTQCASSTY.G
HlSP  VALLKLDSDLSFNSTIKAIELT..N.EIPPEYADAIVSGWGETLVPPPG.IPDQLRSVDVKIIHREKCASRNF.G
AgTP  SLLELEDELTFSDAVQPVGLPKQD.ETVKDGTM.TTVSGWGNTQ.SAAES.NAVLRAANVPTVNQKEC.NKAYSD
                                            ***             **           *

      226       ◆           ◆   ▼                   ◆                         300
PiTP  AGWTITDNMLCSGWLDVGGRDQCQGDSGGPLFHNRIVVGVCSWGLGCADSF..YPGVNARVSRYTAW.IQANA..
CfTP  VGGTITDNMLCSGWLDVGGRDQCQGDSGGPLFHNNVVVGVCSWGQSCA..LARYPGVNARVSRFTAW.IQANA..
MsTP  IGRTVTDNMLCSGWLDVGGRDQCQGDSGGPLYHNGVVVGVCSWGEECA..LARFPGVNARVTRYTSW.ISNNS..
LcTP  YGSQIKPSMICA.Y.TV.GKDSCQGDSGGPLVSGGRLVGVVSWGYGCA..FANYPGVYADVAALRTWVVSAASSV
HlSP  YGSNIKASMICA.Y.AI.GKDSCQGDSGGPLVVNNLLVGVVSWGIDCA..RPSYPGVYVDVSHVRSLIVSNAESI
AgTP  FGG.VTDRMLCAGYQQG.GKDACQGDSGGPLVADGKLVGVVSWGYGCA..QAGYPGVYSRVAVVRDW.VRENSGV
       *      * *       * * *********      *** ***  **      ***   *
```

Figure 3. Predicted amino acid sequence of trypsinogen-like protein from *P. interpunctella* and alignment with serine proteases from five other insect species. PiTP, trypsinogen-like sequence of *P. interpunctella*; CfTP, trypsin of *C. fumiferana*; MsTP, trypsin of *M. sexta*; LcTP, trypsin of *L. cuprina*; HlSP, serine protease of *H. lineatum*; AgTP, trypsin of *A. gambiae*. Functionally important residues are indicated by bold letters and solid triangles (▼) on the top of sequences. Cysteines corresponding to the sites of the predicted disulphide bridges are marked with bold letters and solid diamonds (◆) on the top. Identical residues among all six sequences are indicated with stars (*) at the bottom of sequences. The arrow (↓) indicates predicted signal peptide cleavage site. The arrow (⇓) indicates the N-terminal residues of the active enzymes. Dots (…) represent sequence alignment gaps.

of ≈ 900 bp (Fig. 4), which corresponded in size to the cloned trypsinogen-like protein cDNAs. Northern blot analysis revealed that the expression levels of the trypsinogen-like enzyme mRNA in fourth-instar larvae were similar in the RC688 and HD198 strains.

Southern blot analysis of genomic DNA

Southern analysis was performed on genomic DNA from RC688 and HD198 strains of *P. interpunctella* (Fig. 5). Trypsinogen-like protein cDNA sequences from both RC688 and HD198 strains contained no restriction sites for the three enzymes used, *Pst* I, *Eco*R I and *Hin*d III. The Southern blot revealed three fragments (≈ 1.2, ≈ 3 and ≈ 5.5 kb) in the RC688 strain and five fragments (≈ 1.5, ≈ 2.2, ≈ 3, ≈ 4 and 5.5 kb) in the HD198 strain to be present in *Pst* I + *Eco*R I digests. The trypsinogen-like enzyme cDNA probe hybridized to two fragments (≈ 3 and ≈ 5.5 kb) in both the RC688 and HD198 genomic DNAs which were produced by digestions with *Eco*R I + *Hin*d III restriction enzymes. *Hin*d III + *Pst* I digests showed two hybridized bands (≈ 2.8 and ≈ 6.6 bp) in the RC688 strain

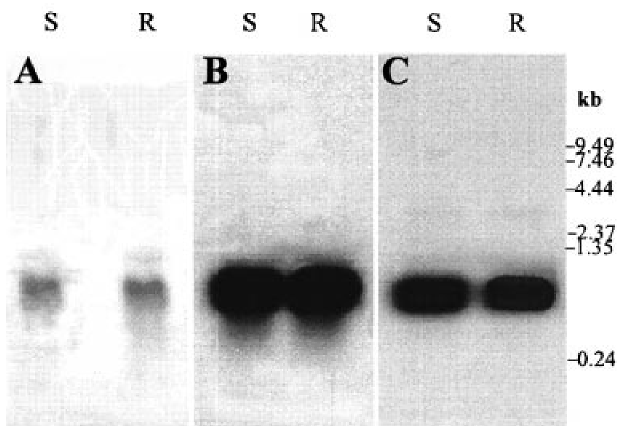

Figure 4. Comparison of trypsinogen-like mRNA expression levels in two *P. interpunctella* strains. S, RC688 strain; R, HD198 strain. (A) Five micrograms of mRNA per lane from each strain was hybridized with an ≈ 300-bp biotin-labelled cDNA probe cloned obtained by PCR using a degenerate primer and a 3′-end vector primer. (B) Three micrograms of mRNA from each strain was loaded in each lane. Messenger RNA was hybridized with a full-length trypsinogen-like cDNA labelled with ^{32}P-dCTP. (C) Same nylon membrane from (B), hybridized with a ribosomal protein S3 cDNA probe.

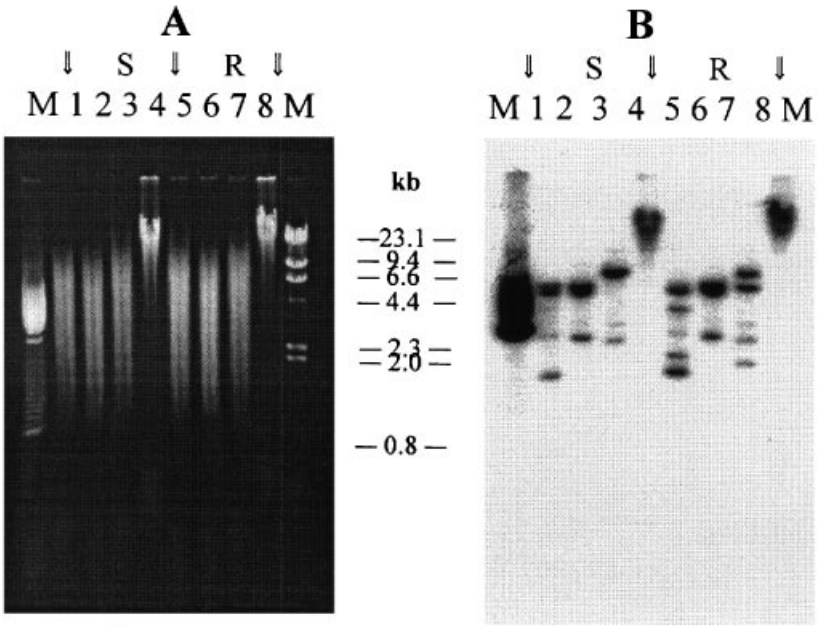

Figure 5. Southern analysis and comparison of genomic DNA from RC688 (lanes 1–4) and HD198 (lanes 5–8) of *P. interpunctella*. S, RC688 strain and R, HD198 strain. (A) Ethidium bromide stained gel; 5 µg/lane of double-digested DNA was separated on 0.8% agarose gel; enzymes used for lanes 1 and 5 = *Pst* I + *Eco*R I, lanes 2 and 6 = *Eco*R I + *Hin*d III, lanes 3 and 7 = *Hin*d III + *Pst* I, lanes 4 and 8 = control; M = DNA marker (left marker was 100-bp ladder from Pharmacia and right marker was high-range ladder from Bio-Rad). (B) Autoradiograph of Southern analysis; lanes are the same as those denoted in (A); digested DNA was transferred to nylon and hybridized with a full-length trypsinogen-like cDNA labelled with ^{32}P-dCTP.

DNA and five bands (≈ 2, ≈ 2.8, ≈ 3.3, ≈ 6.6 and ≈ 7 kb) in the HD198 strain DNA (Fig. 5). Results from the Southern analysis indicated that two additional *Pst*I sites were present in the DNA from the HD198 strain than that from the RC688 strain. These restriction sites were probably located either in introns of the trypsinogen-like protein gene or in untranscribed fragment regions.

Discussion

Midgut cells in most insects synthesize and secrete digestive enzymes in response to the presence of protein in the midgut (Shambaugh, 1954; Felix *et al.*, 1991; Borovsky *et al.*, 1996). It is probable that *P. interpunctella* larvae regulate gut proteinase expression and activity in the same way as most other insects when responding to protein in their food. The HD198 strain has been maintained on artificial diet supplemented with Bt for 110 generations, whereas the RC688 strain has been maintained on a non-Bt diet. Phenotypic proteinase patterns of BApNA-hydrolysing proteinases are the same, however, for HD198 when fed either Bt-treated or untreated diets (B. Oppert, unpublished data).

Serine proteinases are present in the insect alimentary tract primarily to hydrolyse proteins in food. Several insect trypsinogen-like protein cDNAs have been cloned and sequenced (Barillas-Mury *et al.*, 1991; Wang *et al.*, 1993; 1995; Peterson *et al.*, 1994; Noriega *et al.*, 1996). The nucleotide sequences of trypsinogen-like protein cDNAs cloned from a *P. interpunctella* whole-body cDNA library were identical to a cDNA sequence derived from gut RNA. It is likely that trypsinogen-like mRNAs in *P. interpunctella*, like other lepidopteran insects, are expressed predominantly in digestive tissues. Wang *et al.* (1995) demonstrated that trypsin-like enzyme gene expression in *C. fumiferana* was tissue specific and only occurred in digestive tissues. No RNA expression was detected in non-digestive tissues of *C. fumiferana*. Gut-specific expression of digestive proteinases was also observed in the European corn borer, *Ostrinia nubilalis* (Hübner) (Houseman & Chin, 1995) and in the fruit fly, *Drosophila melanogaster* (Matsumoto *et al.*, 1995).

The putative trypsinogen-like protein of *P. interpunctella* is most similar to *C. fumiferana* and *M. sexta* gut trypsin-like enzymes with 68% and 63% identities, respectively. When the amino acid sequence of the *P. interpunctella* trypsinogen-like protein was aligned with those of other insect serine proteinases and subjected to analysis by the CLUSTAL W multiple alignment program (Thompson *et al.*, 1994), the unrooted tree exhibited the *P. interpunctella* protein on the same branch as trypsin-like proteins from *M. sexta* and *C. fumiferana* (Table 1; Reeck *et al.*, 1999).

The mRNAs encoding a trypsinogen-like proteinase were expressed in both strains at similar levels during the fourth instar. Based on the predicted size of the enzyme derived from this gene, we propose that the gene encodes the zymogen for a 25-kDa trypsin-like enzyme, T2, whose expression is invariant among all Indianmeal moth strains tested to date (B. Oppert, unpublished data). Although not implicated in resistance to Bt, this enzyme may be an important target of proteinase inhibitors if it is essential for survival of the insect. Therefore, elucidation of the structure and physiological function of T2 may be valuable information for the development of new biocontrol agents for the Indianmeal moth.

Multiple trypsin-like genes commonly occur in both

mammals and insects (Stevenson *et al.*, 1986; Wang *et al.*, 1995). In lepidopteran insects, three trypsin-encoding cDNAs were isolated from *M. sexta* (Peterson *et al.*, 1994), but only a single, highly expressed trypsin gene has been described in *C. fumiferana* (Wang *et al.*, 1995). In our study, two nearly identical trypsinogen-like cDNAs were cloned and sequenced from RC688 and HD198 strains of *P. interpunctella*. Because specific primers were designed according to the sequence of the trypsinogen-like protein clone that was first isolated and these primers were the only ones used to isolate clones of trypsinogen-like cDNA during the cDNA library screening, the resulting sequences were limited to highly similar trypsinogen-like cDNAs. There may be other less similar trypsinogen-like protein cDNAs present. Although comparable mRNA expression levels were observed in the two strains, genomic organizations deduced by restriction site analysis of trypsinogen-like genes were different. The presence of multiple restriction enzyme fragments hybridizing to a trypsinogen-like enzyme cDNA indicated the possibility of multiple trypsinogen-like protein genes in the *P. interpunctella* genome, although gene mapping and genomic sequencing are required to determine this with certainty.

Our results support the hypothesis that there are no major differences in the 25-kDa trypsin-like proteinases (T2) and their genes isolated from the Bt-susceptible and -resistant strains of the Indianmeal moth examined here. Bt resistance, however, is associated with the absence of a larger 45 kDa trypsin-like proteinase (T1) and an examination of the properties of that enzyme and its gene in *P. interpunctella* will be a focus of future studies.

Experimental procedures

Insect cultures

Plodia interpunctella strain RC688 (Bt-susceptible strain) was collected from farm-stored grain in Riley County, Kansas, and maintained on a diet described by McGaughey & Beeman (1988). A Bt-resistant strain, HD198, was selected from RC688 using Bt ssp. *entomocidus*, strain HD-198 (McGaughey & Johnson, 1992). The resistance level of strain HD198 to Bt ssp. *entomocidus* is 100-fold higher than that of strain RC688 (McGaughey, 1985).

Proteinase activity

Activity blot analysis was performed as previously described (Oppert & Kramer, 1998). Briefly, late fourth-instar larvae were chilled, the posterior and anterior ends removed, and guts were excised and immediately submersed in ice-cold 200 mM Tris, pH 8.0, 20 mM $CaCl_2$ (buffer A) and frozen at –20°C. Fifteen microlitres of pooled gut extracts (5 guts/125 µl buffer) were separated by sodium dodecyl sulphate polyacrylamide electrophoresis (SDS-PAGE) using precast 10–20% Tricine gels (Novex, San Diego, CA) and then electroblotted to nitrocellulose. Nitrocellulose blots were incubated with a substrate solution consisting of *N*-α-benzoyl-L-arginine ρ-nitroanilide (BApNA) (0.5 mg/ml) in buffer A. Released nitroanilide was diazotized by subsequent incubations of 5 min each in 0.1% sodium nitrite in 1 M HCl, 0.5% ammonium sulphamate in 1 M HCl, and 0.05% *N*-(1-naphthyl)-ethylenediamine in 47.5% ethanol.

RNA and poly(A) RNA purification

Two hundred and fifty fourth-instar larvae (approximately 2.5 g wet weight) from each strain were ground in liquid nitrogen. Total RNA was extracted with guanidine thiocyanate solution and precipitated with isopropanol (Titus, 1991). The poly(A) RNA was isolated from the total RNA by chromatography on an oligo(dT)-cellulose column (Gibco BRL Life-Technologies, Gaithersburg, MD).

Construction of cDNA libraries

Double-stranded cDNA was synthesized using 5 µg of poly(A) RNA as a template (ZAP-cDNA synthesis kit, Stratagene, La Jolla, CA), directionally cloned into a UniZAP XR vector phage (Stratagene), and packaged using the ZAP-cDNA Gigapack II Gold cloning system (Stratagene). Approximately 2 and 1.7 million recombinants were obtained for RC688 and HD198 cDNA libraries, respectively.

Development of probes

Lambda DNA of an amplified RC688 cDNA library was prepared using phage precipitation and phenol/chloroform extraction procedures after RNase A and DNase I digestions (Titus, 1991). PCR was carried out with a reverse vector primer T7 and a forward degenerate primer, 5′-TGYCARGGNGAYWSNGGNGGNCCNYT-3′, designed from a highly conserved region (CQGDSGGPL) in both *M. sexta* trypsin and chymotrypsin cDNAs located approximately 250 bp from the 3′ end (Peterson *et al.*, 1994, 1995). PCR-amplified DNA fragments (≈ 300 bp) were cloned into a TA vector (Promega, Madison, WI). Sequences of these clones were obtained through thermal cycle sequencing reactions, the products of which were visualized on silver-stained polyacrylamide gels (Promega, Madison, WI).

cDNA library screening

The cDNA libraries were plated to a density of ≈ 130 plaque-forming units per square centimetre, and plaques were transferred to nylon membranes (MSI, Westboro, MA). The membranes were hybridized at 55°C (LabLogix DNA/RNA Hybridization Kit, LabLogix Inc. Belmont, CA) with a PCR-amplified trypsinogen-like cDNA probe labelled with $\alpha - {}^{32}$P-dCTP (Amersham, Arlington Heights, IL). Nylon membranes were washed at 55°C for 1 h with three changes of 0.2× SSC (sodium chloride/sodium citrate)/0.1% SDS and then exposed to X-ray films. Positive clones from the first library screen were subjected to PCR amplification and analysed for the presence of an insert fragment of the predicted size (≈ 300 bp when amplified by degenerate and T7 priming, ≈ 1 kb when amplified by T3 and T7 priming). Positive clones after PCR screening were plated to a density from which a single plaque could be readily obtained without cross-contamination. PCR amplification was repeated to locate a single plaque with an insert of the predicted size from each clone identified with the ^{32}P-labelled probe. These positive clones were subcloned *in vivo* into pBlue-Script SK(+/–) phagemid (Stratagene, La Jolla, CA). The cDNA inserts were sequenced by primer walking from both

directions using thermal cycle sequencing (Promega, Madison, WI) and an automated sequencer (Applied Biosystem Model 393A) located at the College of Veterinary Medicine, Kansas State University, Manhattan.

Sequence analysis

The Blastx non-redundant program of the National Center for Biotechnology Information Internet server was used to perform similarity searching and retrieval of homologous sequences (Altschul *et al.*, 1990; Gish & States, 1993). The Wisconsin Sequence Analysis Package GCG Unix version 8.1 (Genetics Computer Group, Madison, WI) and sequence analysis tools of the SWISS-PROT Internet server were used to process data of deduced protein sequences.

Gut trypsinogen-like protein cDNA sequence

To verify that trypsinogen-like protein cDNAs cloned from whole-body cDNA libraries represent trypsinogen-like cDNAs from digestive tissues, a RT-PCR was used to amplify a trypsinogen-like cDNA fragment from gut RNA of the RC688 strain. Total RNA was extracted from gut tissues of fourth-instar larvae. Reverse transcription was performed using a SuperScript™ preamplification system for first-strand cDNA synthesis (Gibco BRL Life-Technologies, Gaithersburg, MD) and two specific primers flanking to 5′ and 3′ ends of trypsinogen-like cDNA. The cDNA was amplified using a seminested PCR, and the PCR product (≈ 550 bp) was purified and cloned into a pCR II vector (Invitrogen, San Diego, CA). Sequences were obtained as described and were compared with the trypsinogen-like cDNAs.

Trypsin-like enzyme gene expression in Rc688 and HD198 strains

Trypsin-like enzyme gene expression in fourth-instar larvae of the two strains of *P. interpunctella* was analysed by Northern blotting (Ausubel *et al.*, 1994). Five micrograms of mRNA from each strain was subjected to 1% agarose/formaldehyde gel electrophoresis and transferred to a nylon membrane (MSI, Westboro, MA). An ≈ 300 bp probe was PCR-amplified from identified clones using vector primers T7 and SP6 and was labelled with biotin (Sigma, St Louis, MO). The target RNA was hybridized with the labelled probe at 50°C for approximately 16 h (LabLogix Inc. Belmont, CA). The membrane was developed with a luminescent DNA/RNA detection kit (LabLogix Inc., Belmont, CA) and exposed to X-ray film. To verify that mRNA expression levels was probed with a partial fragment of trypsinogen-like protein cDNA, Northern blot analysis was repeated using a full-length trypsinogen-like protein cDNA as probe. Messenger RNA was purified from another group of larvae and the Northern blotting was conducted with 3 µg of mRNA per lane. The mRNA was transferred to a nylon membrane after gel electrophoresis and hybridized with a full-length trypsinogen-like cDNA probe labelled with $\alpha-{}^{32}$P-dCTP at 55°C for 20 h. The nylon membrane was washed three times and exposed to X-ray film. To evaluate the quantity of RNA loaded in each lane, the trypsinogen-like enzyme cDNA probe was removed from the nylon membrane using a boiling 0.5% SDS solution after the film was processed. The membrane was rehybridized with an α-^{32}P-labelled ribosomal protein S3 cDNA from *M. sexta* as an internal standard, which is highly conserved in sequence among bacteria, yeast, vertebrates, and invertebrates (Jiang *et al.*, 1996).

Southern blot analysis of genomic DNA from RC688 and HD198 strains

Genomic DNA was extracted from fourth-instar larvae using an isolation buffer containing 100 mM Tris-HCl (pH 9), 1% SDS and 100 mM EDTA. Southern analysis was used to search for trypsin-like enzyme gene differences between the *P. interpunctella* strains following procedures of Southern (1975), Wang *et al.* (1995) and Ausubel *et al.* (1994). Three double restriction enzyme digestions, *Pst* I + *Eco*R I, *Eco*R I + *Hin*d III, *Hin*d III + *Pst* I and a control were conducted for each insect strain. In each digestion, 5 µg of genomic DNA was double-digested with 25 units of each enzyme for 5 h at 37°C. Digested DNA was separated using a 0.8% agarose gel, transferred on to a nylon membrane (MSI, Westboro, MA), and hybridized with a full-length trypsinogen-like protein cDNA probe labelled with $\alpha-{}^{32}$P-dCTP at 55°C for 20 h using a DNA/RNA hybridization kit (LabLogix Inc., Belmont, CA). The nylon membrane was washed three times and exposed to X-ray film.

Acknowledgements

The authors wish to thank Haobo Jiang, Yang Wang and Michael R. Kanost of Kansas State University, and James E. Baker and Richard W. Beeman of GMPRC for their technical advice, Donovan E. Johnson and Richard Hammel of USDA, Manhattan, for providing laboratory space and experimental materials, Joel Perez-Mendoza, Sandy Steele, and Michele Zuercher for their assistance in the laboratory. The authors are grateful to Gerald Reeck and Michael Denton of Kansas State University for dendrogram analysis of the amino acid sequences of serine proteinase-like proteins from insects, and also Gerald Reeck and Michael R. Kanost of Kansas State University for reviewing an early version of this manuscript.

References

Altschul, S.F., Gish, W., Miller, W., Myers, E.W. and Lipman, D.J. (1990) Basic local alignment search tool. *J Mol Biol* **215**: 403–410.

Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A. and Struhl, K. (eds) (1994) *Current Protocols in Molecular Biology*. John Wiley & Sons.

Barillas-Mury, C., Graf, R., Hagedorn, H.H. and Wells, M.A. (1991) cDNA and deduced amino acid sequence of a blood meal-induced trypsin from the mosquito, *Aedes aegypti*. *Insect Biochem* **21**: 825–831.

Borovsky, D., Janssen, I., Vanden Broeck, J., Huybrechts, R., Verhaert, P., De Bondt, L., De Bylemans, D. and Loof, A. (1996) Molecular sequencing and modeling of *Neobellieria bullata* trypsin evidence for translational control by *Neobellieria* trypsin-modulating oostatic factor. *Eur J Biochem* **237**: 279–287.

Casu, R.E., Jarmey, J.M., Elvin, C.M. and Eisemann, C.H. (1994) Isolation of a trypsin-like serine protease gene family from the sheep blowfly *Lucilia cuprina*. *Insect Mol Biol* **3**: 159–170.

Felix, C.R., Betschart, B., Billingsley, P.F. and Freyvogel, T.A. (1991) Post-feeding induction of trypsin in the midgut of *Aedes aegypti* L. (Diptera: Culicidae) is separable into two cellular phases. *Insect Biochem* **21**: 197–203.

Gish, W. and States, D.J. (1993) Identification of protein coding regions by database similarity search. *Nat Genet* **3**: 266–272.

Houseman, J. and Chin, P. (1995) Distribution of digestive proteinases in the alimentary tract of the European corn borer *Ostrinia nubilalis* (Lepidoptera: Pyralidae). *Arch Insect Biochem Physiol* **28**: 103–111.

Jiang, H., Wang, Y. and Kanost, M.R. (1996) Primary structure of ribosomal protein S3 and S7 from *Manduca sexta*. *Insect Mol Biol* **5**: 31–38.

Matsumoto, I., Watanabe, H., Abe, K., Arai, S. and Emori, Y. (1995) A putative digestive cysteine proteinase from *Drosophila melanogaster* is predominantly expressed in the embryonic and larval midgut. *Euro J Biochem* **227**: 582–587.

McGaughey, W.H. (1985) Insect resistance to the biological insecticide *Bacillus thuringiensis*. *Science* **229**: 193–195.

McGaughey, W.H. and Beeman, R.W. (1988) Resistance to *Bacillus thuringiensis* in colonies of Indianmeal moth and almond moth (Lepidoptera: Pyralidae). *J Econ Entomol* **81**: 28–33.

McGaughey, W.H. and Johnson, D.E. (1992) Indianmeal moth (Lepidoptera: Pyralidae) resistance to different strains and mixtures of *Bacillus thuringiensis*. *J Econ Entomol* **85**: 1594–1600.

Milne, R.E. and Kaplan, H. (1993) Purification and characterization of a trypsin-like digestive enzyme from spruce budworm (*Choristoneura fumiferana*) responsible for the activation of δ-endotoxin from *Bacillus thuringiensis*. *Insect Biochem Mol Biol* **23**: 663–673.

Moire, N., Bigot, Y., Periquet, G. and Boulard, C. (1994) Sequencing and gene expression of hypodermins A, B, C in larval stages of *Hypoderma lineatum*. *Mol Biochem Parasitol* **66**: 233–240.

Muller, H.M., Crampton, J.M., della Torre, A., Sinden, R. and Crisanti, A. (1993) Members of a trypsin gene family in *Anopheles gambiae* are induced in the gut by blood meal. *EMBO J* **12**: 2891–2900.

Noriega, F.G., Wang, X.Y., Pennington, J.E., Barillas-Mury, C. and Wells, M.A. (1996) Early trypsin, a female-specific midgut protease in *Aedes aegypti*: isolation, amino-terminal sequence determination, and cloning and sequencing of the gene. *Insect Biochem Mol Biol* **26**: 119–126.

Ogiwara, K., Indrasith, L.S., Asano, S. and Hori, H. (1992) Processing of delta-endotoxin from *Bacillus thuringiensis* subsp. kurstaki HD-1 and HD-73 by gut juices of various insect larvae. *J Invertebr Pathol* **60**: 121–126.

Oppert, B. and Kramer, K.J. (1998) Economical proteinase activity blot. *Amer Biotech Lab* **16**: 20.

Oppert, B., Kramer, K.J., Beeman, R.W., Johnson, D. and McGaughey, W.H. (1997) Proteinase-mediated insect resistance to *Bacillus thuringiensis* toxins. *J Biol Chem* **272**: 23473–23476.

Oppert, B., Kramer, K.J., Johnson, D., Upton, S.J. and McGaughey, W.H. (1996) Luminal proteinases from *Plodia interpunctella* and hydrolysis of *Bacillus thuringiensis* CryIAc protoxin. *Insect Biochem Mol Biol* **26**: 571–584.

Peterson, A.M., Barillas-Mury, C.V. and Wells, M.A. (1994) Sequence of three cDNAs encoding an alkaline midgut trypsin from *Manduca sexta*. *Insect Biochem Mol Biol* **24**: 463–471.

Peterson, A.M., Fernando, G.J.P. and Wells, M.A. (1995) Purification, characterization and cDNA sequence of an alkaline chymotrypsin from the midgut of *Manduca sexta*. *Insect Biochem Mol Biol* **25**: 765–774.

Reeck, G., Oppert, B., Denton, M., Kanost, M., Baker, J. and Kramer, K. (1999) Insect proteinases. *Proteases* (Turk, V., ed.), Birkhäuser, Basel, Switzerland. in press.

Shambaugh, G.F. (1954) Protease stimulation by foods in adult *Aedes aegypti* L. *Ohio J Sci* **54**: 151–160.

Southern, E.M. (1975) Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J Mol Biol* **98**: 503–517.

Stevenson, B.J., Hagenbuchle, O. and Wellauer, P.K. (1986) Sequence organization and transcriptional regulation of the mouse elastase II and trypsin genes. *Nucleic Acids Res* **14**: 8307–8331.

Thompson, J.D., Higgins, D.G. and Gibson, T.J. (1994) CLASTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice. *Nucleic Acids Res* **22**: 4673–4680.

Titus, D. (1991) *Promega Protocols and Application Guide*, 2nd edn. Promega Corporation.

Walsh, K.A. and Wilcox, P.E. (1970) Serine proteases *Methods of Enzymology*, Vol. 19 (Colowick, S.P. and Kaplan, N.O., eds). Academic Press, New York.

Wang, S., Magoulas, C. and Hickey, D.A. (1993) Isolation and characterization of a full-length trypsin-encoding cDNA clone from the lepidopteran insect, *Choristoneura fumiferana*. *Gene* **136**: 375–376.

Wang, S., Young, F. and Hickey, D.A. (1995) Genomic organization and expression of a trypsin gene from the spruce budworm, *Choristoneura fumiferana*. *Insect Biochem Mol Biol* **25**: 899–908.